

4.6 Family Background & Demographic Characteristics

This section discusses a variety of demographic, family, and neighborhood factors that potentially affect a young person's labor market attachment and educational attainment. Data supplied by the household informant, the youth, and the parent provide researchers with a wealth of background information about the youth's family members and household environment.

Table 1 overviews the *NLSY97 User's Guide* subtopics available in this section and any global universe restrictions affecting those subtopics. Users should note that certain questions or groups of questions within a subtopic may have additional restrictions. Please consult the NLSY97 questionnaire and codebook for more information on particular questions.

**4.6 Table 1. Family Background & Demographic Characteristics
Subtopics and Universe Restrictions**

<i>NLSY97 User's Guide</i> Subtopic	Round 1 Universe (Age as of 12/31/96)	Rounds 2–4 Universe (Age as of 12/31/96)
4.6.1 Age of Respondent	all ages	all ages
4.6.2 Gender	all ages	all ages
4.6.3 Geographic Indicators	all ages	all ages
4.6.4 Household & Neighborhood Environment	all ages	≤14
4.6.5 Household Composition	all ages	all ages
4.6.6 Nonresident Characteristics	all ages	—
4.6.7 Parent Characteristics		
Background	all ages; parent int. conducted	—
Status	all ages; parent int. conducted	<i>Household Income Update completed</i>
History	all ages; parent int. conducted	—
4.6.8 Race & Ethnicity	all ages	—
4.6.9 Youth History	all ages; parent int. conducted	all ages

Created Variables. In addition to the variables created by CHRR, Child Trends, Inc., an organization involved in the NLSY97 questionnaire design process, has created a number of scales and indexes from several groups of variables described in this section. These scales and indexes are intended to aid researchers in using the various data items relating to environmental and parent characteristics. These variables are described in the following sections:

- Household & Neighborhood Environment
- Parent Characteristics

Although these Child Trends created variables are described only briefly in this guide, interested researchers may obtain a detailed discussion of the creation procedures from NLS User Services. This document, *NLSY97 Codebook Supplement* Appendix 9, also summarizes statistical analyses of the scales and indexes, as well as related data items, performed by Child Trends researchers. (Although considered part of the *Codebook Supplement*, this appendix is physically a separate document.)

4.6.1 Age of Respondent

The NLSY97 collected information on the respondent's age and date of birth in the round 1 *Screener, Household Roster, and Nonresident Roster Questionnaire*. This information was then verified or corrected by the respondent in the initial *Youth Questionnaire* and by the responding parent in the round 1 *Parent Questionnaire*. Date of birth and age were not recollected from respondents in subsequent rounds.

Created Variables. Created variables summarize the information described above, presenting the ages of the respondents as of the interview date in both a regular year format (CV_AGE_INT_DATE) and in a continuous month scheme [CV_AGE(MONTHS)_INT_DATE]. See section 4.4, "Event History," for a more detailed description of the continuous month format. In round 1, the ages of respondents were also calculated as of December 31, 1996 (CV_AGE_12/31/96).

4.6.1 Table 1. Age of NLSY97 Respondents as of 12/31/96

Age (birth year)	Male	Female	Total
12 (1984)	911	860	1771
13 (1983)	934	873	1807
14 (1982)	953	888	1841
15 (1981)	947	927	1874
16 (1980)	854	837	1691
Total	4599	4385	8984

Note: Table based on KEY!SEX and KEY!AGEDOL.

4.6.1 Table 2. Age of NLSY97 Respondents as of Interview Date

Age	Round 1	Round 2	Round 3	Round 4
12	1169	—	—	—
13	1726	105	—	—
14	1858	1597	109	—
15	1877	1676	1664	58
16	1719	1736	1632	1555
17	614	1747	1728	1667
18	21	1416	1622	1657
19	—	109	1387	1606
20	—	—	67	1402
21	—	—	—	136
Total	8984	8386	8209	8081
Noninterview	—	598	775	903

Note: Table based on CV_AGE_INT_DATE.

Data hint →

Researchers may notice that there are two round 1 variables that present the youth's age: KEY!AGE (R05366.) and CV_AGE_INT_DATE (R11941.). The KEY!AGE variable is a symbol generated during the administration of the *Screener, Household Roster, and Nonresident Roster Questionnaire* (round 1). Thus, it initially represented the respondent's age at the screener date. During the youth interview, respondents had an opportunity to correct their age if it had changed since the administration of the screener or was reported incorrectly by the household informant. If a new age was provided, the KEY!AGE variable was updated. However, approximately 400 respondents either had a birthday between the screener date and the youth interview date or had incorrect information collected during the screener but did **not** correct their age information during the interview.

The created age at interview date variable, CV_AGE_INT_DATE, presents the youth's age at the time the round 1 youth questionnaire was administered. This variable was constructed using birth date and interview date information collected during the interview. Because KEY!AGE may not always accurately reflect the respondent's age at the **youth** interview date, researchers may wish to use the created variable for most research purposes. Table 3 below shows the difference in ages between KEY!AGE and the created age variable.

There are also two age variables in the round 2 data, SYMBOL!KEY!AGE (R25162.) and CV_AGE_INT_DATE (R25535.). For almost all respondents, these variables present identical ages, so either can be used. However, two respondents (identification numbers 471 and 476) have inaccurate ages in the symbol variable due to a minor data error. For these two respondents, the created variable contains accurate age information.

In the round 3 data, SYMBOL!KEY!AGE and CV_AGE_INT_DATE present identical information.

The round 4 data has one different age between these two variables: ID number 9022. The CV_AGE_INT_DATE variable has the correct information.

4.6.1 Table 3. Comparison of Round 1 Variables KEY!AGE and CV_AGE_INT_DATE

KEY!AGE	CV_AGE_INT_DATE							Total
	12	13	14	15	16	17	18	
12	1168	63						1231
13	1	1663	79	1				1744
14			1779	79	1			1859
15				1796	93			1889
16				1	1625	91	1	1718
17						523	8	531
18							12	12
Total	1169	1726	1858	1877	1719	614	21	8984

Note: Table based on KEY!AGE (R05366.) and CV_AGE_INT_DATE (R11941.).

User Notes: Because some respondents were first interviewed during the refueling period in March through May 1998, there are respondents in the sample who had already turned 18 by their round 1 interview date. These respondents were still eligible for the sample because they were age 16 or younger as of December 31, 1996. See section 2.2, “Interview Methods,” for more information.

The youth respondent and the responding parent both had an opportunity to verify or correct the respondent’s age as reported in the *Screener, Household Roster, and Nonresident Roster Questionnaire*. However, if the youth and parent did not provide the same age and/or date of birth, no attempt was made to correct this discrepancy. Users should make decisions about these cases on an individual basis. The created variables for age are based on the youth data.

Comparison to Other NLS Surveys: Age data are available for all NLS cohorts. These variables include both the age of the respondents as of a fixed date during the initial survey year and as of the interview date in various years. Date of birth is also available for all cohorts. Consult the *NLS Handbook* or the appropriate cohort’s *User’s Guide* for information about accessing these data.

Survey Instruments: This information was collected in the household roster section of the *Screener, Household Roster, and Nonresident Roster Questionnaire* (round 1). Verification or correction of the information is found in the youth information section (YINF) of the round 1 *Youth Questionnaire* and the PINF section of the *Parent Questionnaire* (round 1).

<i>Related User’s Guide Sections</i>	<i>CD—Main Areas of Interest</i>	<i>CD—Supplemental Area of Interest</i>
4.6.5 Household Composition	Created Variables	Common Variables
4.6.6 Nonresident Characteristics	Demographic Indicators	
4.6.7 Parent Characteristics		

4.6.2 Gender

Information on the respondent's gender was first collected in the *Screener, Household Roster, and Nonresident Roster Questionnaire* (round 1). This information was then verified or corrected by the respondent in the round 1 *Youth Questionnaire* and by the responding parent in the *Parent Questionnaire* (round 1). The interviewer recorded the respondent's gender in the rounds 2–4 interviewer remarks sections (YIR). Table 1 summarizes the composition of the cohort by gender and sample type.

4.6.2 Table 1. NLSY97 Composition by Gender and Sample Type

Sample type	Male	Female	Total
Round 1			
Cross-sectional	3459	3289	6748
Supplemental	1140	1096	2236
Total	4599	4385	8984
Round 2			
Cross-sectional	3213	3066	6279
Supplemental	1070	1037	2107
Total	4283	4103	8386
Round 3			
Cross-sectional	3144	3029	6173
Supplemental	1026	1010	2036
Total	4170	4039	8209
Round 4			
Cross-sectional	3098	2957	6055
Supplemental	1019	1007	2026
Total	4117	3964	8081

Note: Table is based on CV_SAMPLE_TYPE, KEY!SEX, and RNI.

Information on the gender of household members is found in section 4.6.5, “Household Composition.”

Comparison to Other NLS Surveys: Information on gender is available for both the NLSY79 and the Children of the NLSY79. Gender is implicit by membership in the Original Cohorts.

Survey Instruments: This information was collected in the household roster section of the round 1 *Screener, Household Roster, and Nonresident Roster Questionnaire*. Verification or correction of the information is found in the youth information section (YINF) of the round 1 *Youth Questionnaire*, the PINF section of the round 1 *Parent Questionnaire*, and the interviewer remarks section (YIR) of the rounds 2, 3, and 4 *Youth Questionnaires*.

Related User's Guide Sections
 4.6.5 Household Composition
 4.6.6 Nonresident Characteristics

CD—Main Area of Interest
 Demographic Indicators

CD—Supplemental Area of Interest
 Common Variables

4.6.3 Geographic Indicators

The variables in this section are created to describe the respondent's area of residence. The software package Maptitude (V4.2) is used in the creation of the NLSY97 geographic variables. This program assigns latitude and longitude information to respondent addresses. Latitude and longitude are then used to link respondent addresses to standard geographic information such as state, county, and metropolitan statistical area. See the *NLSY97 Geocode Codebook Supplement* for more information about the use of Maptitude.

Main File Created Variables. The NLSY97 main data set includes several created variables that provide general information about the respondent's permanent residence. The main variables indicate whether the respondent lives in an urban or rural area (CV_URBAN-RURAL), whether the respondent lives in a Metropolitan Statistical Area (CV_MSA), and in which Census region the respondent resides (CV_CENSUS_REGION). The states comprising each Census region are listed in Figure 1.

In addition, the same three variables were created in round 1 to describe the respondent's residence as of his or her 12th birthday; another counts the number of residences in which the respondent has lived from his or her 12th birthday until the survey date (all rounds). See section 4.6.9, "Youth History," for more details on these variables.

4.6.3 Figure 1. Number of NLSY97 Respondents and States by Census Region

Census Division	# of Respondents				States		
	Round 1	Round 2 ¹	Round 3 ¹	Round 4 ¹			
Region 1: Northeast	1585	1451	1403	1390	Connecticut Maine Massachusetts	New Hampshire New Jersey New York	Pennsylvania Rhode Island Vermont
Region 2: North Central	2050	1903	1836	1774	Illinois Indiana Iowa Kansas	Michigan Minnesota Missouri Nebraska	North Dakota Ohio South Dakota Wisconsin
Region 3: South	3359	3160	3116	3074	Alabama Arkansas Delaware District of Columbia Florida	Georgia Kentucky Louisiana Maryland Mississippi North Carolina	Oklahoma South Carolina Tennessee Texas Virginia West Virginia
Region 4: West	1990	1846	1834	1808	Alaska Arizona California Colorado Hawaii	Idaho Montana Nevada New Mexico	Oregon Utah Washington Wyoming

Note: Table based on CV_CENSUS_REGION.

¹ Round 2 numbers do not total to 8386, the number of respondents interviewed. Data on geographic location were missing for 4 respondents, and 22 respondents resided outside of the United States. Similarly, the round 3 numbers do not total to 8209 because data were missing for 2 respondents and 18 respondents resided abroad. In round 4, data were missing for 21 respondents and 14 respondents resided abroad, so the numbers do not total to 8081.

Geocode CD-ROM. In addition to all main file data, the NLSY97 Geocode CD provides a variety of statistics for the counties where respondents lived at each interview date. Most of these data are based on the 1994 edition of the U.S. Census Bureau's *County and City Data Book* (the most recent edition available). The first group of geocode variables lists basic demographic information for respondents' counties. These data include land area in square miles; population by race, age, and gender; and birth and death rates. Another variable reports the percent of persons in that county who lived in a different house and/or state in 1990 compared to their residence in 1985, providing information about migration rates for the respondent's area.

Factors that might influence the respondent's education and employment outcomes are the focus of several other geocode variables. These provide the number of serious crimes, households with children, female householders with no spouse present, persons with high school or college degrees, and families below the poverty level. A pair of variables summarizes medical availability for each county, reporting the number of active nonfederal physicians and community hospital beds.

Geocode variables associated with economic and labor force issues include the size of the county's civilian labor force, the percent employed in various industries, and the percent of workers age 16 and over with jobs outside their county of residence. Income variables include per capita money income for the respondent's county, per capita personal income, and median family money income. The unemployment rates for the respondent's metropolitan areas or for the portions of the state not in an MSA (for non-MSA respondents) are also reported.

The final geocode variables involve colleges attended by respondents. Survey staff use information from the Integrated Postsecondary Education Data Systems (IPEDS) to provide users with a code identifying each college attended by the respondent and its location. These codes can be used to associate the NLSY97 respondent's college with various characteristics of the institution contained in the IPEDS database. More information is located in Attachment 102 of the *NLSY97 Geocode Codebook Supplement*.

User Notes: Researchers should be aware that the geocoding procedure changed between rounds 2 and 3 of the NLSY97. All geocode variables from all rounds have been recreated using the new, more accurate method. Regardless of the round of interest, researchers should use data from a round 3 or later Geocode CD. For more information on geocoding procedures, refer to Appendix 5 in the main file *Codebook Supplement* or the introduction to the *Geocode Codebook Supplement* (the latter is distributed only with a Geocode CD).

Due to the detailed nature of the information found on the Geocode CD, use of this dataset is restricted to those meeting confidentiality requirements. The Geocode CD is released only to those who satisfactorily complete the Bureau of Labor Statistics geocode agreement procedure. The application for this geocode agreement is now available online at <<http://www.bls.gov/nls/geocodeapp.htm>>. The round 4 Geocode CD is scheduled for release in August 2002.

Comparison to Other NLS Surveys: Data on the respondent's area of residence are available for all cohorts. For the NLSY79, information on the respondent's region of residence and geographic mobility is available to all users on the main public data file; more detailed information (e.g., state and county of residence) is available on the restricted-use Geocode CD. Data on NLSY79 Children are available through the mother's record. Region of residence and geographic mobility of Original Cohort respondents are provided for most survey years. For more complete information, refer to the *NLS Handbook* or the appropriate cohort's *User's Guide*.

CD—Main Area of Interest: Created Variables

4.6.4 Household & Neighborhood Environment

In the self-administered section of the *Youth Questionnaire*, respondents born in 1982, 1983, or 1984 report on their household environment in the past month. These questions inquire as to whether the respondent usually had a computer available and a quiet place to study at home in the past month. Round 1 also asked if the respondent usually had electricity and heat at home when needed and if a dictionary was usually available at home. Respondents in this age group (12–14 years old) living with a parent-figure or guardian at the survey date are also asked about their household environment in a typical week. In this set of questions, respondents report the number of days in a typical week the housework gets done (round 1 only), their family eats dinner together, their family does something religious together, and their family does something fun together. In round 1, youths also stated the number of days in a typical week that they heard gunshots in their neighborhood.

A separate set of questions in round 1 asked all youths, regardless of age, about events they may have experienced before turning 12 years old. These events include: the respondent's house or apartment was broken into, he or she was the victim of repeated bullying, and he or she saw someone get shot or shot at with a gun. A follow-up asks the respondent to state his or her relationship (e.g., relative, friend, stranger) to the person who was shot or shot at. The same questions, referring to the youth's experiences from age 12 to age 18, were repeated in subsequent rounds for older respondents. The questions are also asked in the youth's first interview after turning 18.

In addition, the survey questions all youths on whether there were gangs in their neighborhood or school at the time of the survey. A gang is defined as a group that hangs out together, wears gang colors or clothes, has set clear boundaries of its territory or turf, and protects its members and turf against other rival gangs through fighting or threats. A separate set of questions asks whether any of the respondent's brothers, sisters, cousins, or friends belongs to a gang. For information on the respondent's participation in a gang, see section 4.5.5, "Crime, Delinquency & Arrest."

Created Variables. Child Trends, Inc., has created several indexes based on the variables described in this section. The following are available:

- (1) Index of Family Routines–Youth Report (FP_YHROUTIN); all rounds. Based on the questions about the number of days per week the youth's family eats together, does housework, does something fun together, and does something religious together, this index indicates the number of days per week spent in routine activities with the family.
- (2) Family/Home Risk Index (FP_ADHRISKI); round 1 only. This index combines responses to a number of questions in the youth and parent interviews, as well as the interviewer remarks section of both the youth and parent questionnaires, to provide an overall assessment of the youth's environment. These variables include the quality of the youth's physical home environment, risk factors in the youth's neighborhood, the availability of study materials in the youth's home, the

religious behavior of the responding parent and the youth's family, the parents' involvement with the youth's schooling, family routines in an average week, time spent watching TV in an average week, the responding parent's physical or mental disabilities, and the parental monitoring and youth-parent relationship scales described in section 4.5.2, "Attitudes."

- Physical Environment Risk Index (FP_ADPENVR1); round 1 only. This index uses a subset of the items from the Family/Home Risk Index to assess risk factors in the youth's physical environment.
- Enriching Environment Index (FP_ADENRCHI); round 1 only. This index uses a second subset of the items from the Family/Home Risk Index to identify opportunities for enriching educational activities in the youth's environment.

Interested researchers should consult *Codebook Supplement* Appendix 9, available from NLS User Services, for more information on the above indexes. In addition to describing the creation procedures for these indexes, the appendix summarizes statistical analyses performed by Child Trends on these variables.

Comparison to Other NLS Surveys: For the NLSY79, information on the respondent's perception of the neighborhood environment is available only for NLSY79 women. Data on both household and neighborhood environments are collected from Children of the NLSY79 age 10 and older (including the Young Adults). For more information, refer to the *NLS Handbook* or the *NLSY79 User's Guide*.

Survey Instruments: These questions are found in the self-administered section (YSAQ) of the *Youth Questionnaire*.

Related User's Guide Section
4.5.5 Crime, Delinquency &
Arrest

CD—Main Areas of Interest
Family Process Measures
Household Characteristics

CD—Supplemental Areas of Interest
Children
Demographic Indicators
Ed Status & Attainment
Marriage & Cohabitation
Parent Current Status

4.6.5 Household Composition

The NLSY97 collects basic demographic information about each member of the respondent's household. In addition, the survey establishes the relationships among household members. Most data were gathered from the household informant in the round 1 household roster and updated by the youth respondent in subsequent rounds.

Data hint ➔

During the interview, the survey program organizes the respondent's answers to questions about household members into a "household roster." This grid of data is the best source of information about household members and should be used instead of the raw data in analyses. Details about the creation and use of the household roster are provided at the end of this section; general information about rosters is available in section 3.2 of this guide.

Screener, Household Roster, and Nonresident Roster Questionnaire (round 1). If a potentially eligible youth lived in the household based on age and date of birth information, the extended screener solicited basic demographic data for each household occupant. In this section, the household informant reported the gender, ethnicity (e.g., Hispanic, Latino, of Spanish origin), and race (e.g., white; black or African American; American Indian, Eskimo, or Aleut; Asian or Pacific Islander; or other) of each occupant.

The household roster, also administered to the household informant, collected further demographic information on household occupants. The survey recorded the highest grade level completed by each household occupant. A follow-up question established the highest degree received by those age 17 and above. This section also collected the marital status of each occupant. Employment information included the number of weeks that each resident above the age of 15 was self-employed or worked for pay at an employee job in 1996. Additional information was collected on the resident's usual number of hours worked per week during that period and his or her current employment status. Finally, the household informant reported the current or most recent occupation of each household resident who either worked at least one week during 1996 or had worked for pay since age 18.

The round 1 household roster further established the relationship of each person in the household to the youth and to each other. Follow-up questions verified the exact relationship; see Figure 1 for definitions of relationships. For example, if the household informant identified an occupant only as a "mother," an additional question asked the household informant if this person was a biological, adoptive, step-, or foster mother. The survey collected the same type of information for a person only identified as a "father." If occupants were listed as half siblings, the interviewer questioned the household informant on whether they shared a biological mother or a biological father. Another set of questions determined if full siblings whose reported birth dates differed by a month or less were multiple births; if they were the same gender, the household informant was asked if they were identical or fraternal twins. In addition, a person listed only

as a “grandmother” was further identified as a maternal, paternal, or social grandmother. The survey solicited similar information for a household occupant listed only as a grandfather, a great-grandmother, or a great-grandfather.

4.6.5 Figure 1. Definitions of Relationships in Household Roster

Biological Relationship: Two people are related by blood or conception and birth. For example, one’s biological father is the same as one’s natural father or the man who made one’s biological mother pregnant.

Step Relationship: Two people are related through a marriage where the husband and/or wife had children with another partner. For example, a stepchild is the biological offspring of one’s spouse (with some other partner). A stepmother is the wife of one’s biological father (if he is not married to one’s biological mother). A stepbrother is the biological son of one’s stepmother who is not the biological son of one’s biological father.

Adoptive Relationship: The permanent legal rights and duties with respect to a child have been transferred from one person or institution to another. The parental rights to an adopted child have been permanently and legally transferred from the birth parents to the adoptive parents. Any other children of the adoptive parents become adopted siblings of the adopted child.

Foster Relationship: Someone assumes a legal and financial obligation for a child but there is no permanent adoptive relationship.

In-law Relationship: Two people are related through marriage. A mother-in-law is the mother of one’s spouse. A son-in-law is the spouse of one’s daughter. A sister-in-law is the sister of one’s spouse or the spouse of one’s brother.

Full Relationship: The youth and his or her siblings share the same biological mother and biological father.

Half Relationship: Siblings share only one common biological parent. Half-siblings have the same biological mother but different biological fathers, or vice versa.

Social Relationship: A person functions in a particular family role but is not biologically related. For example, one’s social grandparent would be someone who functions as a grandparent but is not biologically related.

Source: Interviewer Reference Manual for *Screener, Household Roster, and Nonresident Roster Questionnaire*, 1997.

Parent Questionnaire (round 1). For the round 1 survey, the responding parent provided additional information on the 1996 earnings (self-employment earnings and earnings from an employee job) of each household member older than 14 at the time of the survey. Another question determined the income each household member received from any other sources such as Social Security, pensions, welfare, interest, gifts, etc.

Youth Questionnaire. In rounds 2 through 4, the household information section of this instrument verified characteristics of the respondent’s parents listed in the previous round and asked the youth to describe any times since the last interview when he or she did not live with each parent. Parents who no longer lived with the youth were moved to the nonresident roster (see section 4.6.6, “Nonresident Characteristics”).

The household information section next asked the respondent to review the list of household members from the last interview. If any members moved out of the household or were deceased, this information was recorded and the person was moved to the nonresident roster. The nonresident roster is maintained across rounds so that if anyone moves back into the household, that person can be identified and matched

to their old information by ID number, which remains the same. If anyone from a previous round nonresident roster moved into the household, they were moved onto the household roster but kept the same ID number so that they can be tracked across rounds. Finally, the respondent reported any other new household members, listing their age or birth date and relationship to the respondent. All of the information about current members from the previous round's roster and about new members was then used to create the household roster for the current round.

After the roster was created, the respondent provided additional information about the household members. The questions in this section gathered the gender and race/ethnicity for all new household members, all members transferred from the nonresident roster, and any previous members for whom the information was missing. Marital status, employment status, and highest degree received were recorded for all household members age 16 or older. Current enrollment status was collected for all members age 4 or older; highest grade attended was gathered for all new members and anyone currently enrolled in school. Respondents reported their relationship to any members from the previous round who were not blood relatives or whose relationship was missing. Finally, if the new household member was a stepparent, an adoptive parent, or the partner of the respondent's parent, the date he or she joined the household was recorded.

Created Variables. Using this information, a single created variable (CV_YTH_REL_HH_CURRENT) reports the relationship of the youth to the primary adults in the household at the time of the survey (e.g., both biological parents, biological mother, adoptive parent[s]). In round 1, this variable was also created for various points in the youth's childhood (CV_YTH_REL_HH_AGE_x); see section 4.6.9, "Youth History," for information. Users should be aware that, on the round 1 release, these variables were named CV_HH_REL_BIRTH (for current household composition) and CV_HH_REL_AGE_x. They were subsequently renamed to more accurately reflect the information presented.

Additional created variables identify the household size (CV_HH_SIZE), the number of household members under the age of 6 (CV_HH_UNDER_6), and the number of household members under the age of 18 as of the survey date (CV_HH_UNDER_18). Created variables describing household net worth and total household income are discussed in section 4.6.7, "Parent Characteristics," section 4.8.1, "Assets & Debts," and section 4.8.2, "Income."

User Notes: The classification of a stepparent who has legally adopted a child may differ between the *Screener, Household Roster, and Nonresident Roster Questionnaire* and the other survey instruments. In the former, help screens instructed interviewers to categorize such parent-figures as stepparents rather than adoptive parents. Since the respondent defines this relationship in the self-administered section of the *Youth Questionnaire*, he or she may choose to list this parent-figure as an adoptive parent.

Users should exercise caution when drawing conclusions based on household characteristics using NLSY97 data. The large number of multiple respondent households in the sample may skew the data on certain characteristics if the analysis is performed at the respondent level rather than at the household level. See section 1.3, “NLSY97 Sample,” for more information on multiple respondent households.

Most household information was collected during the administration of the round 1 *Screener, Household Roster, and Nonresident Roster Questionnaire*. This part of the survey also asked questions about any nonresident relatives of the NLSY97-eligible youth, including biological, step-, and adoptive parents and biological children; these data are discussed in detail in section 4.6.6, “Nonresident Characteristics.” The following paragraphs contain a description of the collection and organization of the household roster information. Researchers interested in using these data should first read this discussion. Section 3.2 in this guide contains a general description of NLSY97 rosters.

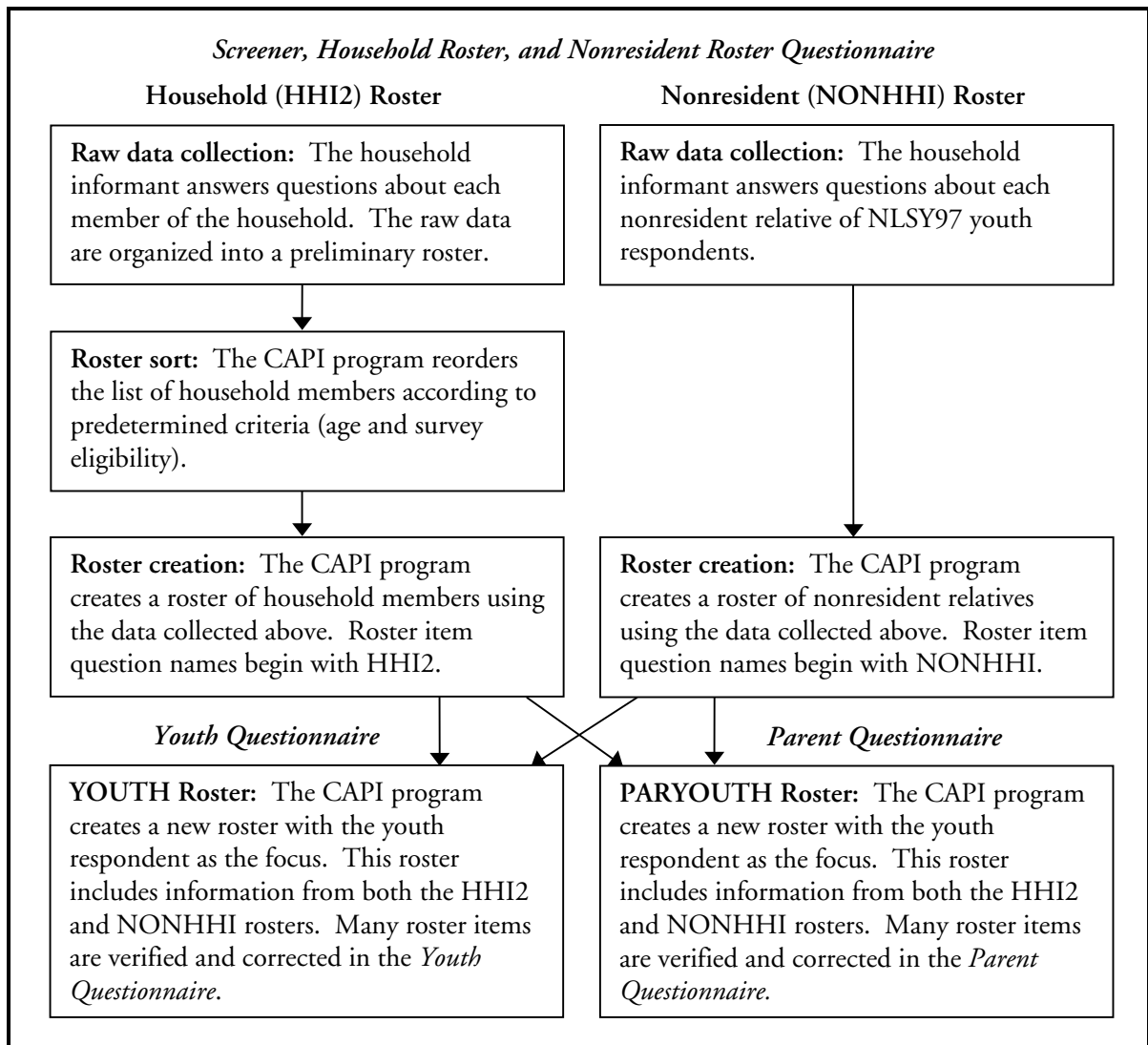
To organize the information about household residents and nonresident relatives, two rosters were created during the administration of the *Screener, Household Roster, and Nonresident Roster Questionnaire*. The first, the household roster, includes information for all current residents of the respondent’s permanent household. For easy identification, all variables on the round 1 household roster were assigned question names that begin with “HHI2_.” The second key roster is the nonresident roster, which presents information about the youth’s nonresident relatives. All question names of nonresident roster variables begin with “NONHHI_.” Note that all household roster variables from subsequent rounds have names that begin with “HHI_” (the “2” is dropped).

Two additional rosters were then created using household and nonresident data provided by the household informant. These youth and parent rosters essentially modify the household and nonresident rosters with the NLSY97 youth respondent as the focus. For example, an item on both new rosters identifies the line where each NLSY97 youth’s biological mother is located. The HHI2 data are not reused because there may be more than one youth and more than one responding parent in a given household, each requiring their own roster (see section 1.3 of this guide for details about multiple respondent households).

The youth roster includes information specific to the NLSY97-eligible youth, as well as some data collected regarding the youth's parents. Many of the roster items are later verified and corrected if necessary during the youth interview. For example, youths are asked if their age as reported by the household informant is correct, and they provide the correct age if the information is inaccurate. All items on the youth roster have question names beginning with "YOUTH_." Similarly, the parent roster is created using *Screener, Household Roster, and Nonresident Roster Questionnaire* data about the youth respondent and the responding parent; it is updated during the parent interview. The question names of parent roster items begin with "PARYOUTH_." The parent roster was created only if a parent interview was conducted.

Figure 2 provides a pictorial representation of how the rosters described above are created and used.

4.6.5 Figure 2. Creation of Round 1 Rosters Based on Screener Data



Users should be aware that much of the information contained in the rosters may appear in the data set more than once. As Figure 2 suggests, data will first be included at the point in the interview when the information was actually collected. For example, screener question SE-28 asked the household informant for the date of birth of each household member. After all the raw data had been gathered, the computer sorted all the answers and created the rosters described above. If there were errors in the original answers and the youth respondent or responding parent provided corrected information, the roster items were often changed to reflect the most up-to-date information. Additionally, because the data were sorted before the creation of the roster, the ID number listed for a person in the screener questions does not necessarily identify the same person as the ID number in the household roster and elsewhere during the interview. To associate screener information with household roster information, researchers must use variables R10978.–R10993., which provide a crosswalk between the two sets of ID numbers. This process can be avoided by using the roster items rather than the raw interview data.

Data hint ➔

Survey staff strongly recommend that researchers use the roster information whenever possible, as it is more accurate and easier to use than the raw data. Roster items are presented as consolidated blocks of data on the public release CD-ROM and can be identified through their unique question names as described above.

In previous CD-ROMs, some questions such as screener question SE-28 were included in the data, but they have now been removed. Survey staff continue to eliminate duplicate sources of information in the data.

Using the household roster. The household roster contains the data described above for each household member and organizes it in a matrix form for use by researchers. A key variable in the household roster is the ID number of the household member (HHI2_ID.xx—the “.xx” indicates that this variable is repeated for each household member, beginning with HHI2_ID.01, HHI2_ID.02, and so on). This variable identifies the line number of the household member on the roster. For example, if the NLSY97-eligible youth is listed first on the roster (that is, has a value of 1 for the variable HHI2_ID.01), then all other HHI2 variables that refer to household member 1 contain information about the youth. If the youth’s father is second on the roster, or has a value of 2 for the variable HHI2_ID.02, then his information is presented in the HHI2 variables referring to household member 2.

Users should be aware that the ID numbers, or line numbers, were assigned by the computer in a specific order. The household informant reported information about household members in no particular order. After the raw data were collected from the household informant, the computer first identified youths eligible for the NLSY97 and put them at the top of the list of residents. If there was more than one eligible youth, the respondents are listed from oldest to youngest. No household has more than five youth respondents, so no youth respondent has a household ID number higher than 5. After listing the

NLSY97-eligible youths, the computer sorted everyone else in the household from oldest to youngest. Therefore, if an older relative such as a grandparent lived in the household, he or she will be listed next, followed in many cases by the youth's parents and then any siblings not eligible for the survey.

The relationship variables on the round 1 household roster provide information about the relationship of every household member to every other household member. For example, consider a household with three members: the respondent (ID number 1), his father (ID number 3), and his grandmother (ID number 2). The resulting relationship variables are depicted in Table 1. In later rounds, relationship variables indicate the relationship of each household member to the respondent but not relationships between household members.

4.6.5 Table 1. Example Structure of the Round 1 HHI2 Relationship Data

Line (ID) number	1	2	3
1	HHI2_REL1.01 (relationship of 1 to 1): identity ¹	HHI2_REL1.02 (relationship of 1 to 2): grandson	HHI2_REL1.03 (relationship of 1 to 3): son
2	HHI2_REL2.01 (relationship of 2 to 1): paternal grandmother	HHI2_REL2.02 (relationship of 2 to 2): identity ¹	HHI2_REL2.03 (relationship of 2 to 3): mother
3	HHI2_REL3.01 (relationship of 3 to 1): father	HHI2_REL3.02 (relationship of 3 to 2): son	HHI2_REL3.03 (relationship of 3 to 3): identity ¹

¹ A code of "identity" for these variables indicates a relationship of "self."

By sorting through the relationship variables, researchers can identify all people in the household with a particular relationship to each other. For example, a user might want to count the number of children of the oldest woman in the household. After identifying which household member is the oldest woman, the researcher can look at each of the relationship variables for that member and see which have a code of son or daughter. If the woman has an ID number of 3, the researcher would write a program that checked the variables for the relationship of member 3 to member 1, member 3 to member 2, and so on. Each member with a code of 49 (daughter) or 50 (son) is a biological child of the woman.

The household roster also includes variables that identify specific types of relationships among household members. For each household member, these variables provide the ID number of that person's biological mother (HHI2_MOMID), biological father (HHI2_DADID), and spouse or partner (HHI2_SPOUSEID or HHI2_PARTNERID) if they also live in the household.

In addition to the set of variables indicating the relationship of each household member to every other person in the household, the round 1 household roster includes a set of variables called HHI2_REL. These variables provide the relationship of each person in the household to the youth respondent,

eliminating the need for the detailed programming described above if the youth respondent is the person of interest. For example, to determine how the youth respondent is related to person 2, researchers can look at the variable HHI2_REL.Y02. In the case presented in table 1 above, this variable would have a value of 4, indicating that household member 2 is the youth respondent's father.

User Notes: Users should note that a number of inconsistencies were discovered during a review of the round 1 relationship data, and the relationship codes in the household roster were substantially revised for the release of the round 2 data. Relationships involving NLSY97 youth respondents were given top priority; some relationships between other household members were updated in the process. Survey staff place greater confidence in the accuracy of codes for relationships involving youth respondents.

Additional corrections were made to relationship and other variables on the household roster for the round 3 Event History release. Researchers should use the most recent data before drawing conclusions in their research. Information about the specific variables affected is provided in the NLSY97 errata notices; contact NLS User Services for details.

Linking the rosters to other data from the same round. Most research requires linking variables from the household and nonresident rosters to other data collected during the parent and youth portions of the survey. This section describes how to identify the youth respondent, responding parent, and household informant, as well as the steps necessary to identify other key relatives of the youth and responding parent.

The youth respondent can be identified by using R05334. (YOUTH_ID.01). This variable provides the line number of the youth respondent on the round 1 household roster. For example, if the value of R05334. is 1, the youth's ID number for the household roster is 1. All information about household member 1 on the roster pertains to the youth. If the value of R05334. is 2, then information about household member 2 pertains to the youth, and so on. As noted above, no NLSY97 youth respondent has an ID number higher than 5, so researchers who just want youth information will only need to examine data for the first five members on the household roster.

The parent roster (PARYOUTH) also contains a variable with the youth ID number on the household roster. However, researchers are advised to use the youth roster variable because it was created during the youth interview.

Identification of the responding parent requires a similar process. Researchers should use variable R07350. (PARYOUTH_PARENTID), which gives the ID number of the parent selected to be the responding

parent at the end of the screener interview. Users should note that the youth roster also contains an ID variable called “ID of R 01 Resp Parent.” However, because this variable is based solely on the screener and does not contain any updated information from the parent questionnaire, researchers are advised not to use this variable to identify the responding parent. (For information about identifying the responding parent’s spouse or partner, refer to section 4.6.7, “Parent Characteristics.”)

Finally, the household informant is fairly easy to identify. Variable R05381. (INFORMANTID) provides the ID number of the informant. As with the parent and youth, this number is the position of the informant on the household roster.

Researchers also may want to identify key relatives of the NLSY97 youth, particularly the youth’s parents, even if they were not respondents to any part of the survey. To facilitate this process, the round 1 data include identification variables that indicate the ID number of a given person on the household and nonresident rosters. Table 2 lists key youth roster ID variables available in the round 1 data set.

4.6.5 Table 2. Round 1 ID Variables for Key Relatives of NLSY97 Respondents

Reference number	Question name (all begin with YOUTH)	Description
R05318.	_ADOPDADID.01	ID # of youth’s resident adoptive father
R05319.	_ADOPMOMID.01	ID # of youth’s resident adoptive mother
R05323.	_DADID.01	ID # of youth’s resident biological father
R05327.	_FOSTDADID.01	ID # of youth’s resident foster father
R05328.	_FOSTMOMID.01	ID # of youth’s resident foster mother
R05336.	_MOMID.01	ID # of youth’s resident biological mother
R05339.	_NONR1ID.01	ID # of youth’s 1 st non-responding resident parent
R05344.	_NONR2ID.01	ID # of youth’s 2 nd non-responding resident parent
R05350.	_NRDADID.01	ID # of youth’s nonresident biological father
R05351.	_NRMOMID.01	ID # of youth’s nonresident biological mother
R05358.	_SPOPARID.01	ID # of youth’s resident spouse or partner
R05359.	_STEPDADID.01	ID # of youth’s resident stepfather
R05360.	_STEPMOMID.01	ID # of youth’s resident stepmother

For instance, the youth respondent answers questions about his or her biological mother. If researchers want to examine the characteristics of the biological mother contained in the household roster or nonresident roster (depending on her residence), they would first look at the MOMID variable. If the mother lives in the household, this variable will have a valid value. For example, a value of 5 means that all the roster variables for household member number 5 contain information about the mother. If there is no positive value for this variable, the next step is to look at the NRMOMID variable to obtain the position

of the biological mother on the nonresident roster. As with the household roster, the value of this variable indicates the ID number of the biological mother on the nonresident roster.

User Notes: Researchers should be aware that the ID variables described above were substantially revised for the release of the round 2 data. The comparable variables on the round 1 CD should **not** be used for analyses. Additionally, researchers may notice that similarly titled variables are contained in the parent roster (PARYOUTH). However, the youth variables (YOUTH) are more accurate because they were adjusted to reflect corrected relationship data. Therefore, researchers are advised to use the youth roster variables to identify the youth's parents.

Linking individuals on the rosters across survey rounds. Many researchers want to examine changes in the youth's household over time. Each round includes the collection of information about members of the youth's household; as in round 1, these data are organized in a roster format. Users should be aware that individuals will not necessarily remain at the same place on the roster in different rounds. For example, a father who had a line ID number of "3" in round 1 might move to position 2 or 4 in round 2. Therefore, each household member is assigned a second, separate identification code, called a unique ID (UID). This unique ID will remain constant across survey rounds, even if household members move to a different place on the roster, so that researchers can identify a given household member in more than one round. For household members in round 1, UIDs are contained in questions HHI2_UID.01–HHI2_UID.16 on the household roster.

As mentioned above, individuals may move between the household roster and the nonresident roster. If such a move occurs, the person keeps the same UID number. Researchers can use the UID number to track individuals as they move in and out of the household. However, users should be aware that no updated information is collected about nonresidents other than biological, step-, or adoptive parents after round 1 (see section 4.6.6, "Nonresident Characteristics," for more information).

The UID number can be used to identify the round in which the household resident was first reported. UIDs that begin with "1" indicate that the person was first reported in round 1 on the household roster. An initial digit of "2" means that the person was first reported on the round 1 nonresident relative roster. If the UID begins with "1998," then the member was first listed on the household roster in round 2, and "1999" indicates a first appearance on the round 3 household roster. For example, a respondent's round 3 household roster might contain the people in Table 3.

4.6.5 Table 3. Examples of UID Codes

Line number	Unique ID	Round first reported
1	102	Round 1 Household Roster
2	202	Round 1 Nonresident Roster
3	199801	Round 2
4	199901	Round 3
5	200001	Round 4

Comparison to Other NLS Surveys: Information on the respondent's household is available for all cohorts for most survey years. Data generally include the age, gender, relationship to the respondent, and educational attainment of all occupants; the enrollment status of those of school age; and the occupation and weeks worked of residents age 14 and older. In the pre-1980 surveys of the Original Cohorts, data were generally collected only for family members living in the respondent's household and not for unrelated household members. For more complete information, refer to the *NLS Handbook* or the appropriate cohort's *User's Guide*.

Survey Instruments: These questions are found in the screener and household roster sections of the round 1 *Screener, Household Roster, and Nonresident Roster Questionnaire* and section P5 of the round 1 *Parent Questionnaire*. In rounds 2–4, they are found in the household information section (YHHI) of the *Youth Questionnaire*.

Related User's Guide Section
4.6.4 Household & Neighborhood
Environment

CD—Main Areas of Interest
Created Variables
Household Characteristics

CD—Supplemental Areas of Interest
Screener Extended
Screener Household

4.6.6 Nonresident Characteristics

In round 1, the household informant provided information about key relatives of the youth who lived outside of the household. In subsequent rounds, the household information section of the *Youth Questionnaire* collected limited information about biological, step-, or adoptive parents who lived outside of the youth's household. This section describes only *Screener* and *Youth Questionnaire* information. The round 1 *Parent Questionnaire* gathered additional background information about nonresident parents; these data are described in section 4.6.7, "Parent Characteristics."

Screener, Household Roster, and Nonresident Roster Questionnaire (round 1). This section describes the data available from the nonresident roster portion of this instrument. It gathered the relationship and status of the youth's nonresident immediate family, including biological, adopted, or stepparents; siblings; spouses; partners; or children. The demographic data collected for these family members are summarized in Table 1 below. Please note that some characteristics, such as gender and marital status, may be inferred from the nonresident's relationship to the youth.

4.6.6 Table 1. Data Collected for Nonresident Relatives in Round 1

Relationship	Gender	Race/ Ethnicity	Age	Marital Status	Highest Grade Level Attained	Highest Degree Attained ¹	Employment Status in 1996
Biological parent		✓	✓	✓	✓	✓	✓
Adoptive parent ²		✓	✓		✓	✓	✓
Stepparent ²					✓	✓	
Full or half sibling ³	✓		✓		✓	✓	
Spouse of youth					✓	✓	✓
Children of youth ⁴	✓				✓	✓	
Parent of youth's children				✓			

¹ This information was only collected if that relative was age 17 or older.

² Information on step- or adoptive parents was collected only if the parent lived with one of the youth's biological parents.

³ Data on full- or half-siblings were collected only if they do not live with either of the youth's biological parents.

⁴ This information was only collected if the youth is age 14 or older.

Data hint ➔

During the interview, the survey program organized the informant's answers to questions about nonresident relatives into a "nonresident roster." This grid of data is the best source of information about these relatives and should be used instead of the raw data in analyses. Details about the creation and use of the nonresident roster are provided at the end of this section; general information about rosters is available in section 3.2.

Additional information about a nonresident biological parent includes the number of miles the youth lived from that parent, the members of that parent's household, and the relationship of each member of that parent's household to the youth. For each nonresident biological child of the youth, the survey collected information on the person the child lived with (e.g., other biological parent, foster parent, adoptive parent, other) and the distance between the child's current residence and the youth's current residence. If nonresident children lived with the other biological parent, the respondent provided additional data on the individuals the other parent lived with (e.g., no other adults, other biological parent's spouse, other biological parent's partner, other biological parent's parents, other relatives).

If the youth's biological parent, full sibling, or child was reported as being deceased, additional information on his or her gender, the year he or she died, and his or her age at death was also gathered.

User Notes: Significant corrections were made to variables on the non-resident roster for the round 3 Event History release. Researchers should use the most recent data before drawing conclusions in their research. Information about the specific variables affected is provided in the NLSY97 errata notices; contact NLS User Services for details.

Youth Questionnaire (rounds 2–4). Although the full set of nonresident questions was not repeated after round 1, subsequent questionnaires collected limited information on biological, step-, and adoptive parents who resided elsewhere. This section focused mainly on whether the youth lived with the nonresident parent at any time between the youth interviews. A series of questions also established whether any nonresident parents shared custody of the youth.

User Notes: After data on nonresident relatives were collected during the administration of the round 1 *Screener, Household Roster, and Nonresident Roster Questionnaire*, the information was organized into a nonresident relative roster. This roster was maintained in subsequent rounds even though no additional questions were asked about most nonresident relatives. The following paragraphs contain a description of the collection and organization of this information and explain how the roster functions in rounds after the initial interview. Before using these data, researchers should first read this discussion. Researchers may also want to refer to the user notes in section 4.6.5, "Household Composition," which contain information helpful in using the nonresident roster, and to section 3.2, which contains general information about NLSY97 rosters.

To organize the information about household residents and nonresident relatives, two rosters were created during the administration of the *Screener, Household Roster, and Nonresident Roster Questionnaire*. The

first, the household roster, included information for all residents of the respondent's permanent household at the time of the round 1 interview. For easy identification, all variables on the household roster were assigned question names that begin with "HHI2_." The second key roster is the nonresident roster, which presents information about the youth's nonresident relatives. All question names of nonresident roster variables begin with "NONHHI_."

Two additional rosters, described in detail in section 4.6.5, were then created using screener data provided by the household informant, modified with the youth respondent as the focus. The YOUTH roster was used during the youth interview, and the PARYOUTH roster was used during the parent interview (if a parent interview was conducted). Figure 2 in section 4.6.5 provides a pictorial representation of how the rosters described above are created and used.

Users should be aware that much of the information contained in the rosters actually appears in the data set more than once, first at the point in the interview when the raw data were actually collected and then in the roster. Survey staff strongly recommend that researchers use the roster information whenever possible, as it is more accurate and easier to use than the raw data. Roster items can be identified through their unique question names as described above.

Using the nonresident roster. The nonresident roster is similar to—but somewhat simpler than—the household roster. Although the roster is not sorted in any particular manner, each person about whom information is collected appears at the same point in the roster for each item. Thus, researchers can easily identify the key characteristics of a given nonresident relative. Different information was collected about nonresident relatives depending on their relationship to the youth, as depicted in Table 1 above.

Researchers often want to identify the respondent's biological mother and father so that they can use the information collected in the nonresident roster and connect it to data collected in other sections of the survey. To facilitate this process, the round 1 data include identification variables that indicate the ID numbers of the nonresident mother (R05351.) and father (R05350.) on the roster. For example, if the youth's biological mother does not reside in the youth's household, the nonresident mother ID variable provides the position of the biological mother on the nonresident roster. If the value of this variable is 3, then all nonresident roster variables referring to person 3 contain information about the youth's mother.

Data hint ➔

Users should note that these variables were substantially revised for the release of the round 2 data and again for the round 3 CD-ROM; they should **not** use the variables from the round 1 CD-ROM in analyses.

User Notes: Although questions about nonresident relatives were not asked in rounds 2–4, the nonresident roster remains part of the data for these rounds. This roster does contain updated information about the respondent’s nonresident biological, step-, or adoptive parents. In addition, anyone who was previously a member of the youth’s household but who left that household is moved to the nonresident roster. This permits the unique ID code (UID) of that person to be maintained across rounds. If the person later returns to the respondent’s household, he or she will be moved back onto the household roster. The UID code can also be used to track nonresident parents across rounds if their information is updated.

Comparison to Other NLS Surveys: For the Children of the NLSY79, data are available about the education, occupation, and geographic residence of the child’s nonresident biological father. Two surveys of the NLSY79, two surveys of the Mature Women, and six surveys of the Young Women have collected demographic information on the respondent’s nonresident siblings. For more complete information, refer to the *NLS Handbook* or the appropriate cohort’s *User’s Guide*.

Survey Instruments: Questions are found in the nonresident roster section of the round 1 *Screener*, *Household Roster*, and *Nonresident Roster Questionnaire*.

<i>Related User’s Guide Sections</i>	<i>CD—Main Area of Interest</i>	<i>CD—Supplemental Areas of Interest</i>
4.6.5 Household Composition	Non-Res Characteristics	Parent Current Status
4.6.7 Parent Characteristics		Parent Retrospective
4.9 Marital History & Fertility		Screener Non-Resident

4.6.7 Parent Characteristics

This section details information collected from the responding parent in the round 1 *Parent Questionnaire*, as well as data gathered in the *Youth Questionnaire* and *Household Income Update* in subsequent rounds. First, several types of background information provided by the responding parent are described. Data on the status of NLSY97 respondents' parent-figures at the interview date are then detailed. The third subsection describes the retrospective information collected about the responding parent.

Background

This subsection of the *Parent Questionnaire* (round 1) first established the responding parent's origin or descent, birth date, and place of birth. If the parent was born in the United States or a U.S. territory, a follow-up question asked for the state or territory of birth. Parents not born in the United States were questioned on their country of birth and the year they first came to live in the United States.

User Notes: Researchers should note that the commonwealth of Puerto Rico is listed as both a country (P2-002) and as a U.S. territory (P2-007). In addition, Quebec is listed as a country separate from Canada. The United Kingdom of Great Britain and Northern Ireland is not listed; respondents specify England, Northern Ireland, Scotland, or Wales.

The survey also collected information about the responding parent's number of siblings, the religion in which he or she was raised, and the year of birth and highest grade level completed by his or her parents. Further questions focused on whether this parent lived with both biological parents while growing up, and if not, the parent he or she did live with, the reason he or she did not live with both parents, and the ages that he or she lived apart from one or both parents. The type of community in which the responding parent lived when he or she was 14 was also gathered. The responding parent reported this same information for his or her current spouse or partner. An additional question determined the number of times the responding parent's current spouse or partner had been legally married. If the NLSY97 respondent had had contact with a nonresident biological parent since the age of 10, the same set of questions was asked about that nonresident parent.

Status

The responding parent was questioned about his or her present employment status, marital status, and highest grade completed, as well as whether any language other than English was spoken at home. These data were also provided for the responding parent's spouse or partner. Additional questions about religion were addressed to parents of youths born in 1982, 1983, and 1984. In the self-administered section of the questionnaire, these parents responded to the following statements about their religiosity:

- I don't need religion to have good values

- The Bible/Koran/Torah/religious teachings should be obeyed in every situation
- I often ask God to help me make decisions
- God has nothing to do with what happens to me personally
- I pray more than once per day

Created Variables. A scale created by Child Trends, Inc., is based on the religious beliefs and practices questions in round 1 and indicates the religiosity of the responding parent (FP_PPRELIG). See *Codebook Supplement Appendix 9*, available from NLS User Services, for more information.

Attitudes: In the self-administered section of the *Parent Questionnaire* (round 1), responding parents of youths born in 1982, 1983, and 1984 expressed their level of agreement with a series of statements about their attitude toward themselves. The same questions were asked of the youth respondent, as described in section 4.5.2, “Attitudes.” Data on the responding parent’s attitude towards his or her partner or spouse and the youth’s view of their relationship are also described in section 4.5.2.

Health: The NLSY97 collected information on the responding parent’s and his or her current spouse’s/partner’s general health. Data were gathered on the presence of any long-term health problem that limited the type or amount of employment possible. The parent then reported any serious long-standing medical condition (e.g., severe heart problem, cancer, potentially life-threatening disease) since the birth of his or her oldest child. The parent’s height and weight were also collected. If the NLSY97 respondent was not currently living with his or her biological parents but had been in contact with them since age 10, the responding parent was asked about the height and weight of those parents. In multiple respondent households, these questions were asked if the oldest respondent who was a child of the biological parents had had contact since age 10.

Income and Assets: In round 1, information was gathered on the responding parent’s earnings from a job—including military service—and from a farm, business, or professional practice during the 1996 calendar year. Similar data were collected for his or her current spouse or partner. Additional questions gathered information about the combined income received from other sources by the responding parent and the spouse or partner during 1996. Sources listed were interest or dividend income, including any amount that has been reinvested or credited; child support payments; and other income. Finally, data were collected on amounts received from various government programs in 1996.

For a more comprehensive picture of the family’s financial situation, the responding parent was also asked about assets and debts. The parent first stated whether the family owned or rented their dwelling (e.g., ranch or farm, mobile home, house or apartment). After determining the percentage of the dwelling and the land owned by the family, a follow-up question asked about the present value of these holdings. The

survey then established whether there was a mortgage, land contract, or any other type of loan that used the property as collateral, such as a second mortgage or a home equity loan, and the amount still owed. If the family neither rented nor owned the dwelling, the living situation was determined (e.g., housing part of job compensation, temporary living arrangements while house is under repair).

Additional asset information included the amount that would be received (less any outstanding debts) if any business partnership/professional practice or any real estate, such as a second home or land, was sold. The responding parent also stated whether he or she owned each of the assets listed in Figure 1; a follow-up question ascertained the total dollar value of each asset reported. Finally, to balance information on assets, the responding parent reported the amount still owed on any vehicles, any educational loans for a child, and other debt including credit cards or bank loans.

4.6.7 Figure 1. Assets Held by the Responding Parent

Bills and bonds including CDs, government savings bonds, or treasury bills
Checking, savings, or money market accounts
Educational IRA accounts or other prepaid tuition savings accounts
Other assets including money owed by others, life insurance policies, and precious metals
Pension or retirement savings
Shares in publicly-held corporations or mutual funds
Vehicles

<p><i>User Notes:</i> Researchers using income and asset data should be aware of survey practices with respect to item nonresponse and topcoding high asset values. To reduce the impact of item nonresponse and capture at least some information, responding parents who don't know or refuse to provide the exact value for an income or asset question are asked to select the appropriate range from a predefined list. Topcoding of parent and household income information uses the same system as youth income variables, as discussed in the introduction to section 4.8 of this guide.</p>
--

Created Variables. Several created variables summarize the household's financial situation. First, the net worth of the household is calculated by subtracting total debts from total assets. If the youth was independent, this calculation is based on information provided by the youth and is reported in the variable CV_HH_NET_WORTH_Y; if the youth was not independent and a parent was interviewed, the calculation is based on parent data and the results reported in CV_HH_NET_WORTH_P. The second variable, CV_INCOME_GROSS_YR, provides gross household income in the previous year. Although only one variable is created, the data for the variable, like net worth, could be taken from the youth questionnaire if the respondent was independent and from the parent questionnaire if the youth was not

independent. The variable CV_HH_INCOME_SOURCE identifies whether the income variable was created from youth or parent information. Finally, a poverty status variable (CV_HH_POV_RATIO) creates a ratio comparing the gross household income variable to the federal poverty level for the previous year, taking household size into account. After round 1, all household income and net worth information are based on data provided by the youth respondent (if he or she is independent). See section 4.8.1, “Assets & Debts,” and 4.8.2, “Income,” for details.

Household Income Update. The surveys following round 1 include this form, administered to one of the respondent’s parents, which is designed to gather basic income information concerning the respondent’s parent and his or her spouse/partner in the absence of a detailed parent questionnaire. All respondents who live with a parent at the interview date are eligible to have this questionnaire administered. It collects the parent’s total pre-tax income from wages, salaries, commissions, and tips during the past calendar year; the same data for the parent’s spouse or partner; and the total pre-tax amount of any other income received (i.e., farm or business income, inheritances, child support, government programs). In round 2, parents of 7,601 respondents answered at least one question from the *Household Income Update*; parents of 5,488 respondents answered at least one question in round 3; and 5,225 parents of respondents answered at least one question in round 4. The administration of this one-page paper instrument is described in section 2.2 of this guide.

User Notes: The *Household Income Update* was administered only once in households with more than one NLSY97 respondent. In round 2, the information provided by the parent was attached to the records of all the NLSY97 respondents who lived in the round 1 household. This means, for example, that if one respondent had left the household to live with a different parent between rounds 1 and 2 but the siblings remained in the original household, that respondent would have the same *Household Income Update* information that was provided by the parent in the original household. In round 3 and subsequent rounds, the instrument reflects the youth’s actual household at the time of that survey.

Youth Questionnaire. In rounds 2, 3, and 4, this instrument included basic questions about the income of the respondent’s parents. These separately referred to the youth’s biological mother, biological father, nonbiological mother, and nonbiological father as applicable. The youth was asked for that parent’s total income from a job, farm, business, or professional partnership. This information, along with the parent’s responses to the *Household Income Update*, may allow researchers to compare a youth’s perception of parental income with the amount the parent reports.

History

In the round 1 parent calendar section, retrospective data were collected on the parent's employment and marriage histories and history of participation in government programs for low-income households (e.g., AFDC or ADC, SSI, food stamps).

Employment: The round 1 NLSY97 collected a history of the responding parent's spells of employment since the parent's 18th birthday or the youth's date of birth, whichever was earlier. In this context, a spell was defined as three or more months of employment without a break of six or more months, regardless of employer. After establishing the start and stop dates for each spell, the responding parent reported the usual number of hours worked per week during each spell. The responding parent also provided work history information for each spouse who lived with the oldest NLSY97 youth during his or her marriage to the responding parent. The survey collected information on whether the spouse was working at the time of the marriage or for a period of three consecutive months or more during the marriage. If these conditions were met, the responding parent provided information on the spouse's spells of employment and the average number of hours per week worked during each spell.

Marriage: Questions on the responding parent's marital history collected information on the length of each of the responding parent's marriages. In addition, the survey gathered data on any changes in the marital status (e.g., legal separation) during each marriage.

Program Participation: A series of questions recorded the number of years during the previous five years that a responding parent participated in various government programs targeting low income households (e.g., AFDC or ADC, SSI, WIC).

Created Variables. Although no questions in the round 1 survey directly collected a fertility history from the responding parent, two created variables based on information from the household and nonresident relative rosters relate to this part of the parent's personal history. These created variables provide the age of the youth's biological mother when she gave birth to her first child (CV_BIO_MOM_AGE_CHILD1) and the mother's age when she gave birth to the youth respondent (CV_BIO_MOM_AGE_YOUTH).

User Notes: Users should exercise caution when drawing conclusions based on parent characteristics using NLSY97 data. The large number of multiple respondent households in the sample may skew the data on certain characteristics if the analysis is performed at the respondent level rather than at the parent level. See section 1.3, "NLSY97 Sample," for more information on multiple respondent households.

Basic demographic information about the responding parent and his or her spouse or partner is also available in the household roster. The content and structure of the household roster are discussed in detail in section 4.6.5, “Household Composition.” The current user note describes how to identify the responding parent’s spouse or partner so that household roster information may be linked to other data about the spouse or partner.

To find the responding parent’s spouse or partner on the household roster, researchers must use the HHI2 relationship variables. Researchers should begin this process by determining the ID number of the responding parent (R07350.). The value of this variable indicates the position of the responding parent on the roster. The next step is to look at the household relationship variables for the responding parent. Codes of 1 (wife), 2 (husband), and 69 (partner) indicate that the household member is the responding parent’s spouse or partner. For example, if the responding parent is number 3, the user would examine the variables for the relationship of member 3 to member 1, member 3 to member 2, member 3 to member 4, and so on. If the variable for member 4 had a code of 2, then member 4 would be the responding parent’s husband.

Comparison to Other NLS Surveys: Data provided by the respondent about his or her parents are available for all cohorts. For the NLSY79, information includes the parents’ birthplace, educational attainment, occupation, age, and life status. Information about the mothers of NLSY79 Children is quite detailed, as the mothers are main respondents for the NLSY79; basic demographic data on the fathers are available as well. Original Cohort respondents have provided information about the country of birth and life status of their parents, as well as the educational attainment and occupation of their parents during their teenage years. More recently, the Mature and Young Women also provided information about their parents’ health and income and about transfers of time and money to and from their parents. For more complete information, refer to the *NLS Handbook* or the appropriate cohort’s *User’s Guide*.

Survey Instruments: These questions are found in sections P2 through P6 of the round 1 *Parent Questionnaire*. The rounds 2–4 questions about parent income are located in the income (YINC) section of the *Youth Questionnaire* and in the *Household Income Update*, a brief questionnaire administered to one of the respondent’s parents.

Related User’s Guide Sections
 4.5.2 Attitudes
 4.6.5 Household Composition
 4.6.6 Nonresident Characteristics

CD—Main Areas of Interest
 Created Variables
 Family Process Measures
 Parent Background
 Parent Current Status
 Parent Family Background
 Parent Retrospective

CD—Supplemental Areas of Interest
 Attitudes
 Household Characteristics
 Income
 Non-Res Characteristics

4.6.8 Race & Ethnicity

Data on the respondent's race and ethnicity were collected in the *Screener, Household Roster, and Nonresident Roster Questionnaire* (round 1) and were based on the household informant's identification. Using the household roster variables, the survey program created KEY!RACE, which describes the respondent's race, and KEY!ETHNICITY, which identifies respondents of Hispanic origin. These variables can be combined to create a single race/ethnicity variable; however, there are a number of missing observations. Researchers may prefer to use a created variable, KEY!RACE_ETHNICITY, described in the user notes below. Table 1 in section 1.3 summarizes the racial/ethnic composition of the sample.

Created Variables. A round 1 created variable (CV_CITIZENSHIP) provides the respondent's U.S. citizenship status based on his or her parents' residence at the time of the respondent's birth. The variable KEY!RACE_ETHNICITY indicates respondents' race and Hispanic ethnicity.

User Notes: In the supplemental PSUs, the information from the *Screener, Household Roster, and Nonresident Roster Questionnaire* was used to determine whether a youth was eligible for inclusion in the black and Hispanic oversample. No non-black/non-Hispanic youths from the supplemental PSUs were included in the survey. For more information on sampling procedures, refer to section 2.1, "Sample Design & Screening Process."

To simplify the race/ethnicity identification process, survey staff created a single combined variable, KEY!RACE_ETHNICITY (R14826.). This variable is based on KEY!RACE, KEY!ETHNICITY, household roster information from the HHI2 variables, and biological parent race/ethnicity information. All respondents are classified as Hispanic, black, non-black/non-Hispanic, or mixed race/non-Hispanic; there are no missing values. While a respondent can be of Hispanic ethnicity and still be of any race, Hispanic ethnicity was given priority in the creation of this variable. Users who wish to identify, for example, blacks of Hispanic origin must create their own variable from the screener information.

Although the questionnaire contains response categories for white; black or African-American; American Indian, Eskimo, or Aleut; and Asian or Pacific Islander races, some household informants refused to select one race and identified household members only as "mixed race." This category was added in the codebook.

Comparison to Other NLS Surveys: Race is available for all cohorts; ethnicity is available for all cohorts except the Older Men and Young Men. Users should be aware that coding categories for race and

ethnicity have varied among cohorts and over time. For more complete information, refer to the *NLS Handbook* or the appropriate cohort's *User's Guide*.

Survey Instruments: This information was collected in the extended screener section of the round 1 *Screener, Household Roster, and Nonresident Roster Questionnaire*.

Related User's Guide Sections

4.6.5 Household Composition

4.6.6 Nonresident Characteristics

4.6.7 Parent Characteristics

CD—Main Areas of Interest

Created Variables

Demographic Indicators

CD—Supplemental Areas of Interest

Household Characteristics

Non-Res Characteristics

Parent Background

4.6.9 Youth History

Parent Questionnaire (round 1). In the youth history section, the responding parent first reported the NLSY97 respondent's housing situation since his or her birth. If the youth lived with anyone else before the responding parent, the survey collected a history of living situations. This history recorded the relationship of each adult the youth lived with and in what month and year the situation changed. All round 1 responding parents were asked whether the youth had ever lived apart from them for three months or more (since the youth first lived with them). For each spell apart, the parent stated the place where or the person with whom the youth lived, whether the arrangement was still in effect, and the year the arrangement changed.

The survey also considered residential moves (not necessarily related to a change in living situation). These questions recorded the total number of different places the youth lived prior to his or her 12th birthday and since his or her 12th birthday.

For each NLSY97 youth not living with both biological parents, the survey asked whether the responding parent holds legal responsibility or legal custody for the youth. If only the youth's biological mother was listed on the birth certificate, follow-up questions asked whether the biological father had ever been legally identified (e.g., blood test, court ruling, signed legal document). For adopted youths, data were collected on whether the youth ever lived with his or her biological mother or father.

Screener, Household Roster, and Nonresident Roster Questionnaire (round 1). Information included the distance the youth lived from his or her biological parents (if they were alive) or the date that the youth's biological parents died.

Youth Questionnaire. In rounds 2 through 4, the respondent answered questions about residential moves to a different city, county, or state since the date of last interview. These data were used to update the total number of residences variable described below. These surveys also collected the periods the youth lived with each parent between interviews.

Created Variables. A single created variable (CV_TTL_RESIDENCES) calculates the total number of different residences in which the youth has lived since the age of 12. A series of round 1 variables, CV_YTH_REL_HH_AGE_x, reports the relationship of the youth to the primary adults in the household (e.g., both biological parents, biological mother, adoptive parent[s]) at ages 2, 6, and 12. Users should be aware that, on the round 1 release, these variables were named CV_HH_REL_AGE_x; they were renamed for subsequent releases to more accurately reflect the content of the variable. A similar variable, CV_YTH_REL_HH_CURRENT (formerly CV_HH_REL_BIRTH), provides the same information as of the interview date for each survey round. Finally, a series of round 1 created variables describing the

youth's geographic residence at age 12 indicates in which Census region the youth resided (CV_CENSUS_REGION_AGE_12), whether the youth lived in an MSA (CV_MSA_AGE_12), and whether the youth lived in an urban or rural area (CV_URBAN-RURAL_AGE_12). For more information about geographic variables, see section 4.6.3, "Geographic Indicators."

User Notes: Researchers should be aware that the geocoding procedure changed between rounds 3 and 4 of the NLSY97. All geocode variables from all rounds have been recreated using the new, more accurate method. Regardless of the round of interest, researchers should use data from a round 4 Geocode CD. For more information on geocoding procedures, refer to Appendix 5 in the main file *Codebook Supplement* or the introduction to the *Geocode Codebook Supplement* (the latter is distributed only with a Geocode CD).

Comparison to Other NLS Surveys: The *Childhood Residence Calendar*, a supplement to the 1988 NLSY79 survey, provided detailed retrospective data on the residential history of each respondent from birth to age 18. For more information, refer to the *NLS Handbook* or the *NLSY79 User's Guide*.

Survey Instruments: Questions are found in section PC8 of the round 1 *Parent Questionnaire* and in the nonresident roster section of the round 1 *Screener, Household Roster, and Nonresident Roster Questionnaire*. Residential move updates are collected in the youth information section (YINF) of subsequent youth questionnaires.

Related User's Guide Sections

4.2.4 School Experience
4.5.2 Attitudes
4.6.3 Geographic Indicators

CD—Main Areas of Interest

Created Variables
Geographic Indicators